IN-SILICO STUDY OF NON-SPECIFIC EGFR INHIBITORS FOR TARGETING EGFRvIII IN GLIOBLASTOMA

A thesis Submitted in Partial Fulfilment

of the Requirements for the Degree of

Bachelor of Technology

in

Biomedical Engineering

Submitted by

Rajlaxmi Choudhury

Under the Guidance of

Dr. B.P.Nayak

Department of Biotechnology and Medical Engineering

National Institute of Technology, Rourkela, India

2012

Department of Biotechnology and Medical Engineering

National Institute of Technology

Rourkela

CERTIFICATE

This is to certify that the thesis entitles **"*In-silico* Study of Non-Specific EGFR Inhibitors for Targeting EGFRvIII in Glioblastoma"** submitted by **Ms. Rajlaxmi Choudhury** in partial fulfilment of the requirements for the degree of Bachelor of Technology in Biomedical Engineering in the department of Biotechnology and Medical Engineering, National Institute of Technology, Rourkela is an authentic research work carried out by him under my supervision and guidance.

To the best of my knowledge, the matter embodied in this report has not been submitted earlier to any university/ Institute for the award of any Degree or Diploma.

Date :

Place : Rourkela

Dr. Bibhukalyan PrasadNayak

Assisstant Professor

Department of Biotechnology and Medical Engineering

National Institute of technology, Rourkela-769008

Contents

ACKNOWLEDGEMENT

I take this opportunity to express my hearty gratitude and profound regards to my thesis supervisor, **Dr.B.P.Nayak** , Assistant Professor, Department of Biotechnology and Medical Engineering, National Institute of Technology, Rourkela, for whose constant guidance and encouragement led to the successful completion of my B.Tech thesis. It is for his esteem guidance and noble supervision which resulted in the materialization of this project. I would also thank him for helping me to infer my mistakes objectively and inspiring me to improve upon it , thus inculcating a scientific temperament and keeping my interest alive in the subject as well as for being approachable at all the times.

I would also extend my sincere thanks to Mr. Satya Teja, M.Tech student for providing me knowledgeable insight regarding Bioinformatics tools which were of much relevance for my project work .

I am also thankful to all the lab-members working under the supervision of Dr.B.P.Nayak for their suggestions and ideas which resulted in several modifications in my data and helped me set up a better presentation.

Finally, I am also grateful to all my batch-mates for keeping up an intellectual environment in the department and their support and unity which kept us patient during difficult times.

Rajlaxmi Choudhury

Department of Biotechnology and Medical Engineering

List of Abbreviations

1. EGFR : Epidermal Growth Factor Receptor
2. GBM : Glioblastoma Multiforme
3. PDB : Protein Data Bank
4. EGF : Epdermal Growth Factor
5. TGF : Transforming Growth Factor
6. DNA : Deoxy-ribo- Nuclein Acid
7. CNS :Central Nervous System
8. PDGF : Platelet Derived Growth Factor
9. VEGF : vascular endothelial growth fator
10. ADT : AutoDock Tools
11. MGL Tools :
12. ACD : advanced chemistry development

List of Figures

List of Tables

ABSTRACT

Glioblastoma (GBM) is known to be an extremely aggressive and malignant form of anaplastic arythrocytoma affecting the central nervous system in humans. Due to its worse prognosis leading to high mortality rate over the years , there is an urgent need to deviate from conventional surgical and radiation therapeutic treatments and delve into molecular targeted drug therapy techniques. Epiderma Growth Factor Receptor(EGFR) and its mutant version of EGFRvIII are known to be overexpressed case of GBM resulting in high rate of proliferation and failure in cell cycle arrest. EGFRvIII is expressed in about 62% of GBM cases and are self-activating receptors independent of ligands. This project is undertaken for the study of potential inhibitors for EGFRvIII mutant using bioinformatics tool. Thirteen of commercially available EGFR inhibitors were docked against EGFRvIII protein and stability of the docked results was based on their free binding energy values. The result was compared with that of Gefitinib and Erlotinib, the established inhibitors of EGFR. PD153035, Lavendustin A and BIBX-1328BS were found to be the most suitable inhibitor for EGFRvIII , which can be proposed for wet-lab testing as molecular targeted drug therapy for GBM. Further, few structural modifications on two of these ligands were performed and *in-silico* docking analysis was repeated to check for more potential alternatives which can prove to be favourable inhibitors.

Keywords : GBM (Glioblastoma Multiforme), EGFRvIII (1I8I PubMed ID), *in-silico* protein-ligand docking.

CHAPTER 1

INTRODUCTION

1. INTRODUCTION

Among the cancers affecting the brain, Glioblastoma Multiforme (GBM) is the most commonly diagnosed. It is also established to be the most dangerous and aggressive form of primary neoplasm detected in central nervous system of humans [1] .GBM arises from the glial tissue of the brain (which include astrocytes, oligodendrocytes and ependymal cells). It has been studied that approximately 60% of the primary brain tumour detected are gliomas [2]. Gliomas are the neoplasms present in the CNS possessing different histological characteristics. Based on the similarities in the histological features between the malignant and the tumour cells, gliomas are broadly categorised into ependymomas, oligodendrogliomas and astrocytomas [3]. Astrocytomas are more prevalent in humans and GBM is its most common and aggressive form [4]. Unfortunately, GBM is more common in adult human males and has rather severe effect on brain than the spinal cord [5].

Generally, cancer consists of cells which continue to proliferate or divide when they are supposed to undergo apoptosis or cell cycle arrest. In most cancers, including GBMs, these kinds of molecular abnormalities are very prevalent. Now the signal for proliferation is specifically provided by a growth factor or ligand. Few important examples are TGF-alpha, platelet derived growth factor (PDGF) , epidermal growth factor(EGF) , vascular endothelial growth factor(VEGF) . These ligands interact with cells by the medium of receptors like EGF receptors, VEGF receptors and PDGF receptors. Activation of receptors is associated with cellular processes of mitosis or invasion into surrounding sites by upstream signal transduction cascades. But in case of cancer cells, these pathways show abnormalities in several levels. In 92% of astrocytoma, EGFR is overexpressed [6] and EGFRvIII, which is a mutant receptor capable of activation independent of any ligand, is expressed in up to 62% of GBMs [7].

Certain receptors or molecular markers that are found overexpressed on malignant cells but are nearly absent in normal cells, act as potential drug targets. Hence, EGFRvIII is considered to be a novel target for pathological techniques. Studies suggest that EGFRvIII is 24% to 65% overexpressed in GBM patients [8]. This mutated version is formed from EGFR by the genetic loss of its 270 amino acids.

GBM continues to remain associated with precarious prognosis [9][10]. While its insensitivity toward conventional chemotherapy treatments remains and the results of drug targeting due to the blood-brain barrier are more or less unsatisfactory [11][12], there is an urgent need for some results in the molecular therapeutic domains. This study adopts a cheminformatics based drug design approach to obtain few potential inhibitors against GBM from 16 non-specific EGFR inhibitors available for commercial usage. Comparative molecular docking experiments were undertaken between the ligands and EGFRvIII. Based on their binding energies and other properties, the most potential drug-like molecules were obtained. Further docking experiments were performed

The motive behind this study was to narrow down the most probable effective commercially available inhibitors against EGFRvIII. This stays an attempt to contribute to the molecular therapeutic studies being carried out for the treatment and an adept prognosis of GBM.

CHAPTER 2

GLIOBLASTOMA MULTIFORME

2. GLIOBLASTOMA MULTIFORME

As it is known, cancers of brain are the consequence of abnormal growth of brain cells. From its origin, brain cancer can be divided broadly into two categories. First are the primary brain tumours. These mainly involve primary brain cells like gliomas, meningiomas, pituitary adenomas, vestibular schwannomas, primary CNS lymphomas, and medulloblastomas. Sometimes the site of cancers might be exclusive of brain but the tumour made of cancerous cells spread into brain by the bloodstream. These are known as metastatic brain tumour or secondary brain tumour.

Glioma is a type of tumour that starts in the brain or spine and 60% of the primary brain tumour detected is gliomas .It derives its name from its association with the glial cells. These can be classified by cell types, grade and location. Gliomas resemble certain histological features with specific cell types , even though they need not necessarily originate from them. Based on this similarity gliomas are categorised as ependymomas (from ependyman cells), astrocytomas (from astrocytes) and oligodendrogliomas (from oligodendrocytes). Among these, astrocytomas are known to be most prevalent in humans. GBM is one of its most aggressive forms. Glioblastomas are also important brain tumour in canines and hence, research is being carried out to use this as a model for developing treatments in humans. GBM involves glial cells and accounts for 52% of all functional tissue brain tumour and 20% of all intracranial tumours.

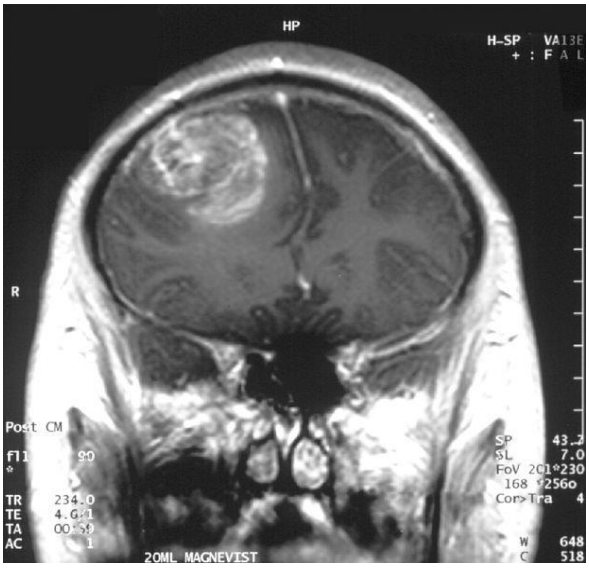


figure 2. 1Coronal MRI with contrast of a GBM WHO grade IV in a 15-year-old male.

Based on the pathologic evaluation of the tumour, WHO (World health organisation) classified gliomas into four grade types:

Grade I: this stage is referred when the tissue is benign and in the least advance stage. Best prognosis can be provided during this phase.

Grade II: this stage is referred when the gliomas are well differentiated (not anaplastic). Even though they are not benign, there are chances of better prognosis of patients.

Grade III: this stage consists of anaplastic gliomas , which are highly malignant. Carrying out a prognosis is extremely difficult.

Grade IV: the disease exists in its most advance stage and the prognosis is worse in this stage.

Tentorium is a membrane present in the brain which seperates the cerebrum from the cerebellum. Based on the presence of the tumour above or below the tentorium, it is known as supratentorial and infratentorial respectively. Supratentorial gliomas persist mostly in adults (70%) and infraperitonial gliomas are majorly found in children.

GBM has several regions of point mutations which establishes self-activating signal transduction pathways of the tyrosine kinase receptors [13] . EGFRvIII is one such mutation of EGFR when it loses 270 of its amino acids. EGFRvIII is overexpressed in about 62% of GBM cases studies. Due to these certain abnormalities GBM is seen to be resistant to conventional therapeutic treatments. GBM possess the inherent nature to diffuse topographically. This renders surgeries very unsatisfactory. The variable location of the tumour cell within the brain results in the inability to completely resect the tumour. Tumour cells has a tendency to migrate through brain parenchyma , get collected just below the pial margin and surround neurons and vessels to migrate through whit matter tracks [14].

Despite various multimodality treatment consisting of open craniotomy with surgical resection of maximum amount of tumour possible followed by chemoradiotherapy and symptomatic managements with corticosteroids, prognosis is extremely poor with a median survival time of approximately 14 months [15]. Most of the drug-targeted therapy remains unsuccessful as the blood brain barrier prevents the drug to reach their respective targets. During the last decade Temozolomide has come up chemotherapeutic drug that is able to cross the blood-brain barrier. Its effectiveness was studied in a large clinical trial consisting of 575 patients completely randomised. Comparative study was carried between patients treated by radiotherapy versus radiation plus temozolomide chemotherapy. Results showed that group receiving temozolomide had median life of 14.6 months as opposed to 12.1 months for group receiving radiation without temozolomide [16]. Although systematic chemotherapy benefited with improved results, an answer for long term disease control is yet to be devised.

CHAPTER 3

EPIDERMAL GROWTH FACTOR RECEPTORS

3. EPIDERMAL GROWTH FACTORS RECEPTOR

The epidermal growth factor receptor is the cell surface receptor member of EGF family protein ligands. It consist of four closely related tyrosine kinase receptors: EGFR(ErbB1),HER2/c-neu(ErbB2),HER3(ErbB3) and HER4(ErbB4). The receptor is present on the cell surface and is activated by the specific ligands like EGF and TGF-alpha. It functions by homo-dimerization with other family member which in turn activates its intrinsic intracellular protein-tyrosine kinase activity. The autophosphorylation of several tyrosine residues in EGFR activates signalling by several downstream proteins in signal transduction cascades, for example MAPK and JNK pathways. This leads to DNA synthesis, cell migration ,adhesion and proliferation. Any mutation that leads to EGFR over expression is directly associated with GBM. This mutation leads to constant activation resulting in uncontrolled cell division.

3.1 EGFRvIII (EPIDERMAL GROWTH FACTOR RECEPTOR Viii)

EGFRvIII is a mutated variation of EGFR which often overexpressed in case of patients suffering from GBM. EGFR losses 270 of its amino acid residues to turn into this particular mutated version. EGFRvIII is capable of self-activation even in the absence of any ligand which might results in abnormal proliferation, metastasis and inhibition of cell death leading to a formation of an aggressive tumour.

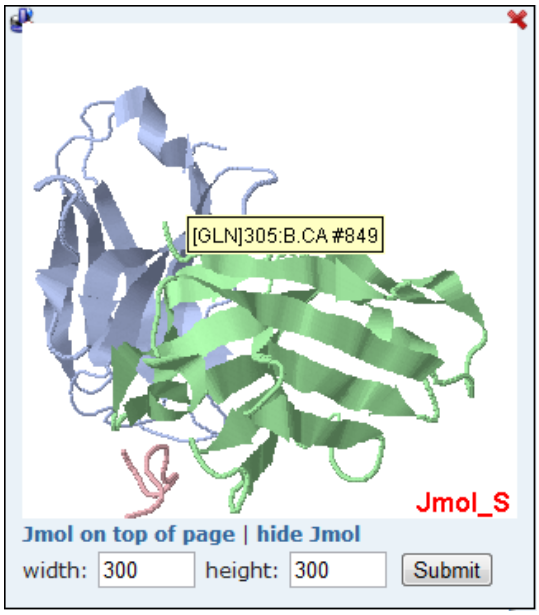


Figure 3. 1EGFRvIII mutant version. Courtesy PDB. Image is viewed using Jmol[19]

The PUBMED ID of EGFRvIII is **1I8I**. EGFRvIII has in total 3 chains. These chains are represented by 3 sequence unique entities.

Below are the information about this protein obtained from PDB:

CHAIN A: EGFR Antibody MRISCFV light chain

This is the first polymeric chain consisting of 107

residues. It is light polymeric chain type. The PDB secondary structure of the protein is given below

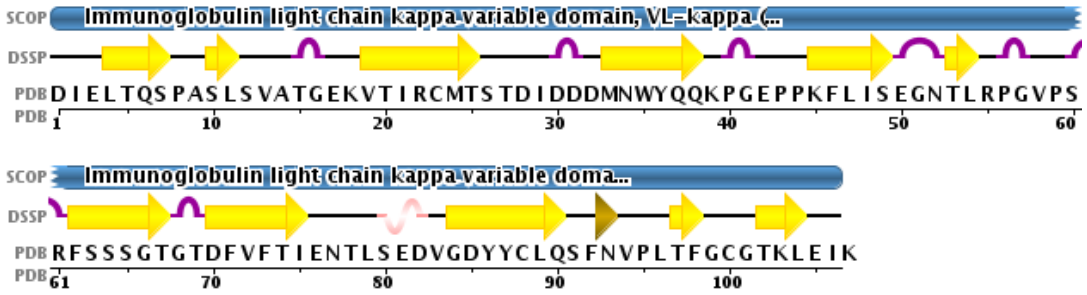


Figure 3. 2Secondary structure of Chain A of EGFR vIII. Source PDB

The secondary structure is 2% helical with 1 helices and 3 residues and 47% beta sheet with 12 strands and 51 residues.

CHAIN B : EGFR Antibody MRISCFV heavy chain

This is the second polymeric chain consisting of 124 residues . It is heavy polymeric chain type. The PDB secondary structure of the protein is given below

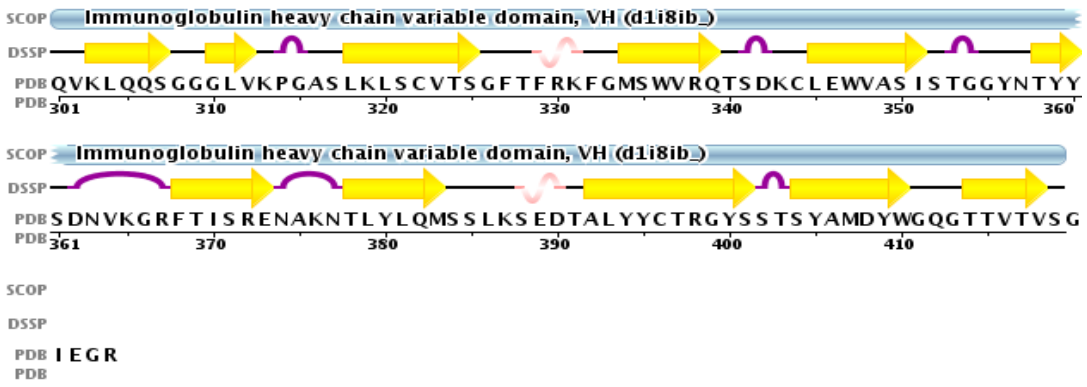


Figure 3. 3Secondary structure of Chain B of EGFRvIII.Source PDB

The secondary structure is 4%helical with 2 helices and 6 residues and 53% beta sheet structure with 11 strands and 66 residues

CHAIN C : EGFRvIII peptide antigen

Figure 3. 4Secondary structure of Chain of EGFRvIII. Source PDB

This is the third polymeric chain of the protein. It contains 12 residues. The PDB secondary structure of the protein is given below

The secondary chain has 8% beta sheet structure with 1 strand and 1 residue.

CHAPTER 4

TOOLS FOR *IN-SILICO* STUDY

4. TOOLS FOR IN-SILICO STUDIES

4.1 PROTEIN DATA BANK

The PDB (protein data bank) is the single worldwide archive of structural data of Biological macromolecules, established in Brookhaven National Laboratories (BNL) in 1971. It contains Structural information of the macromolecules determined by X-ray Crystallography, NMR methods etc.

4.2 PUBCHEM

The PubChem compound database contains validated chemical depiction information provided to describe substances in PubChem Substance. Structures stored within PubChem compounds are pre-clustered and cross-referenced by identity and similarity groups.

4.3 UCSF Chimera

Chimera is developed by the Resource for Biocomputing, Visualization, and Informatics (RBVI) at the University of California,San Francisco, a program for interactive visualization and analysis of molecular structures.it helps us to obtain the structure either directly from the PDB site or through many other formats, to localize the ligand that is in the active site and display the docked structure. Secondary structure identification can be done, as it can show and hide ribbons. The secondary structure can oobserve motifs of the protein and also create a molecular surface around the protein, then colour the surface according to different properties of the amino acids. Studying the protein-ligans interaction canbe done by identifying the residues within 5.0A, and then identify the residues involved in the binding and label the residue names and types. Chimera shows the hydrogen bond interaction between proteins and ligands.

4.3.1 Energy Minimizations

Before energy calculations can be performed, it is necessary to correct structural inconsistencies, add hydrogen, and associate atoms with force field parameters. Clicking Minimize dismisses the dialog (unless the option to Keep dialog up after Minimize is checked) and may call Dock Prep to perform several tasks toprepare structure. Dock Prep may in turn call Add H and Add charge. Each of the tasks is a check-bx option that can be turned off independently if already done or deemed unnecessary. Minimization routines are performed by NMTK, which in included with Chimera.

4.4 AUTODOCK 4.2

(Link: http://autodock.scripps.edu/)

AutoDock is an automated procedure or predicting the interaction of ligands with biomacromolecular targets. The motivation fo r this work arises from problems in the design of bioactive compounds, and in particular the field of computer-aided drug design. Progress in biomolecular x-ray crystallography allows accessto important protein and nucleic acid structures. These structures could be bioactive agents in the control of animal and plant diseases. The study of precise interaction of such agents or candidate molecules with their targets is important in the developmental process. The goal of docking tools is the is to assist researchers in the determination of biomolecular compexes.

In any docking scheme, two comflictin grequirements must be balanced: the desire for a robust and accurate procedure, and the desire to keep the computational demands at a reasonable level. The ideal procedure would find the global minimum in the interaction energy between between the substrate and the target protein, exploring all available degrees of freedom (DOF) for the system. AutoDock combines two methods to achieve these goals: rapid grd-based energy evaluation and efficient search of torsional freedom[20].

4.5 ACD/ChemSketch version 12.0

(link : http://www.acdlabs.com/chemsketch/)

ACD/ChemSketch is an integrated software package from Advanced Chemistry Development, Inc., developed to help chemists quickly and easily draw chemical structures, reactions, and schematic diagrams, calculate chemical properties, and design professional reports and presentations.

ACD/ChemSketch includes:

- Structure mode for drawing chemical structures and calculating their properties (for more information, refer to Section 3).
- Draw mode or text and graphics processing (for more information, refer to Section 4).
- Additional modules that extend the ACD/ChemSketch possibilities (most of them should be purchased separately).

4.6 Q-SiteFinder

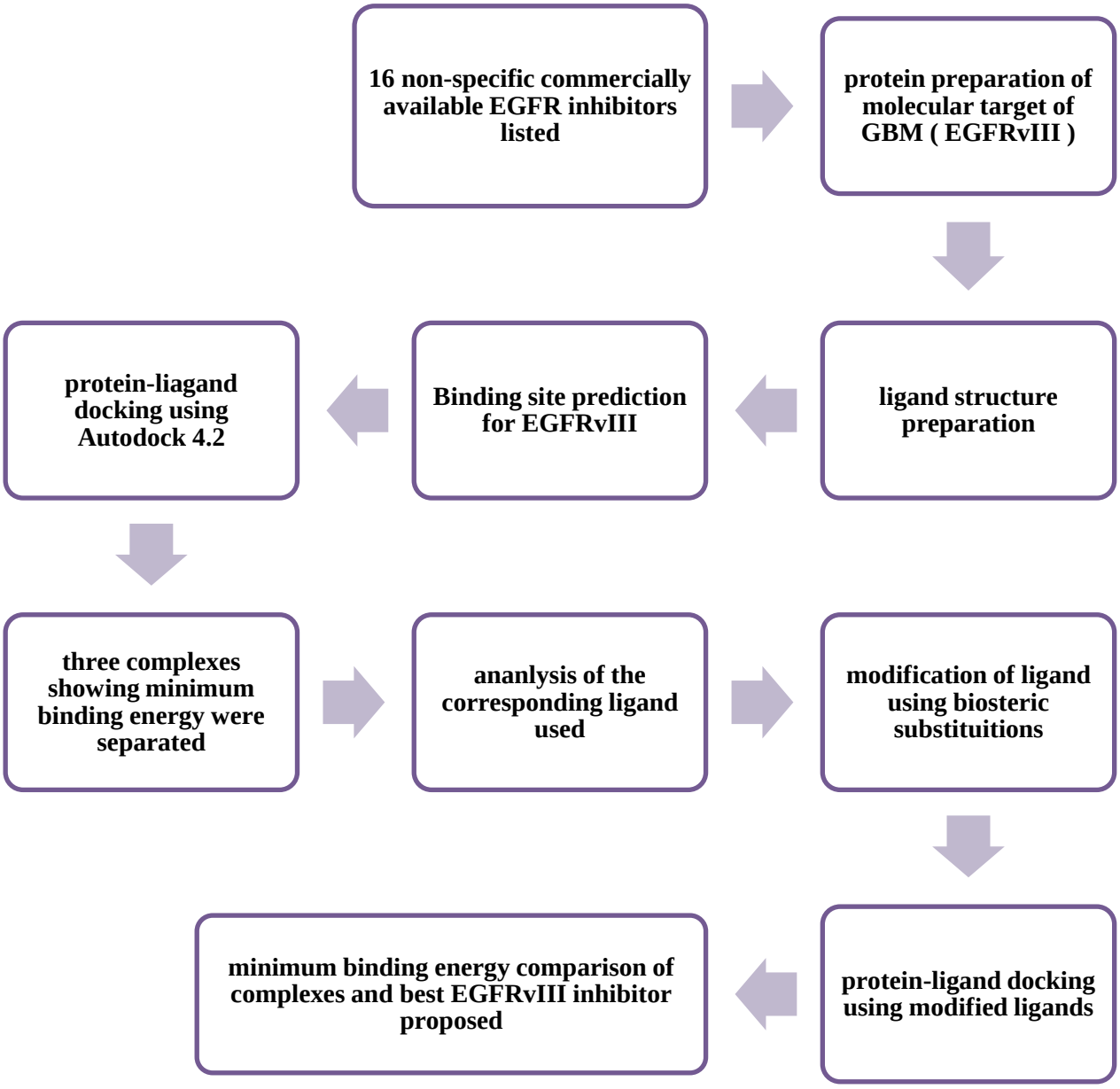
Q-SiteFinder is a new method of ligand binding site prediction. It works by binding hydrophobic (CH3) probes to the protein, and finding clusters of probes with the most favourable binding energy. These clusters are placed in rank order of the likelihood of being a binding site according to the sum total binding energies for each cluster.

It also gives the coordinates of the binding-box around selected sites. Binding box coordinates can be used as input for a docking program to ensure that ligands are only docked onto into the binding site.

CHAPTER 5

METHODOLOGY

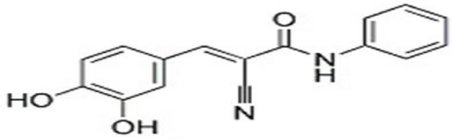
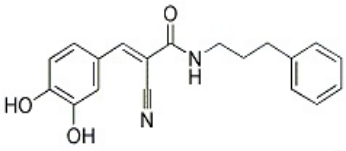
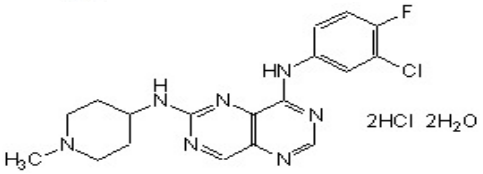
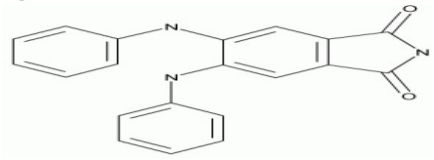
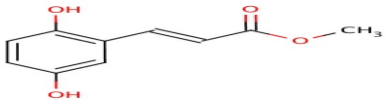
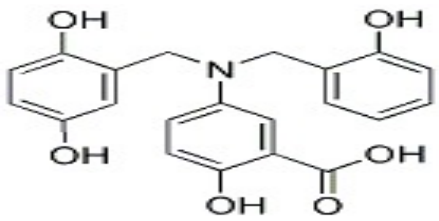
5. METHODOLOGY



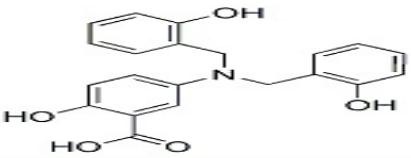
5.1 Ligand structure preparation

The 16 non-specific commercially available EGFR inhibitors were listed. Its commercial usage was confirmed from a private retailer ChemCruz™ Biochemicals , Santa Cruz Biotechnology Inc. PubChem was used to download their chemical structures and other details and was viewed and converted into PDB file using Chimera. Given below are the details of 16 inhibitors whose docking with EGFRvIII protein was analysed

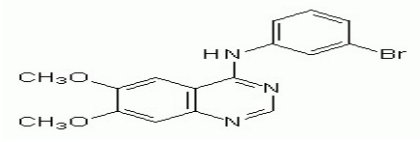
Table 5. 1LIST OF POTENTIAL EGFR INHIBITORS

DETAILS OF LIGAND	CHEMICAL STRUCTURE
1. AG494 PubChem compound ID – 2055 Molecular wt .- 280.28 Molecular formula – $C_{16}H_{12}N_2O_3$	
2. AG555 Pubchem compound ID – 5328770 Molecular wt.- 322.36 Molecular formula- $C_{19}H_{18}N_2O_3$	
3. BIBX -1382BS Pubchem compound ID – 6918505 Molecular wt.- 496.80 Molecular formula- $C_{18}H_{19}CLFN_7.2HCL.2H_20$	
4. DAPH Pubchem compound ID – 1697 Molecular wt.- 329.4 Molecular formula- $C_{20}H_{15}N_30_2$ E-value : - 239.32	
5. **ERBSTAIN ANALOG** Pubchem compound ID – 5353609 Molecular wt.- 194.0 Molecular formula- $C_{10}H_{10}O_4$	
6. LAVENDUSTIN A Pubchem compound ID – 3894 Molecular wt.- 381.38 Molecular formula- $C_{21}H_{19}NO_6$ E-value : -240.72	

7. LAVENDUSTIN B
Pubchem compound ID – 3895
Molecular wt.- 365.4
Molecular formula- $C_{21}H_{19}NO_5$

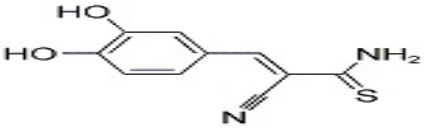


8. **PD 153035**
Pubchem compound ID – 4705
Molecular wt.- 360.21
Molecular formula- $C_{16}H_{14}BrN_3O_2$

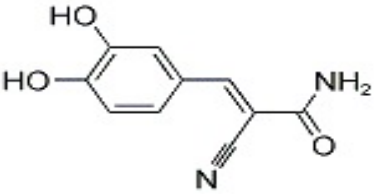


9. RG 13022
Pubchem compound ID – 5468216
Molecular wt.- 266.3
Molecular formula- $C_{16}H_{14}N_2O_2$

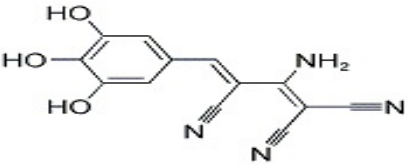
10. Tyrophostin 47
Pubchem compound ID – 5485187
Molecular wt.- 220.2
Molecular formula- $C_{10}H_8N_2O_2S$



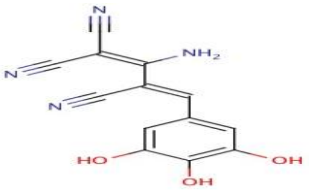
11. TYROPHOSTIN AG99
Pubchem compound ID – 5328567
Molecular wt.- 204.18
Molecular formula- $C_{10}H_8N_2O$



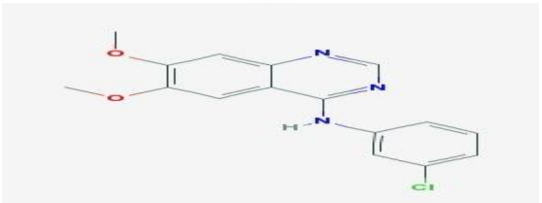
12. TYROPHOSTIN 51
Pubchem compound ID – 5328807
Molecular wt.- 268.2
Molecular formula- $C_{13}H_8N_4O_3$



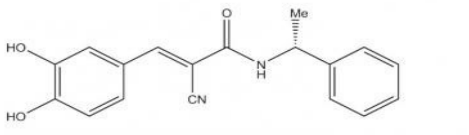
13. TYROPHOSTIN AG 183
Pubchem compound ID – 5328807
Molecular wt.- 268.2
Molecular formula- $C_{13}H_8N_4O_3$



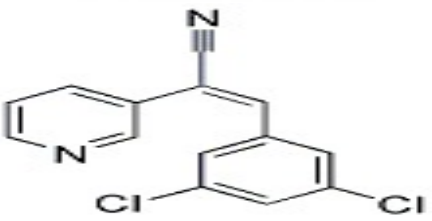
14. TYROPHOSTIN AG1478
Pubchem compound ID – 2051
Molecular wt.- 315.76
Molecular formula- $C_{16}H_{14}ClN_3O_2$



15. TYROPHOSTIN B44(-)
Pubchem compound ID – 5328772
Molecular wt.- 308.3
Molecular formula- $C_{18}H_{16}N_2O_3$



16. TYROPHOSTIN RG 14620
Pubchem compound ID – 5926218
Molecular wt.- 275.14
Molecular formula- $C_{14}H_8Cl_2N_2$



5.2 Binding site prediction for EGFRvIII

The most probable ligand binding site of 1I8I was determined using the Q-SiteFinder version2.3[18] as the active sites for EGFRvIII was not well defined in PDB. The SiteMap predicts a suitable binding site by the following three steps :

- Initially a grid is assigned and all the points are grouped into sets depending on several criterias used to define sites
- The corresponding sites are then mapped on another grid and visual files of the maps are produced
- Finally, all the properties are evaluated and the different sites are written in maestro-readable form. Every step is compiled by an impact job and hence the best site is considered for usage in docking studies.

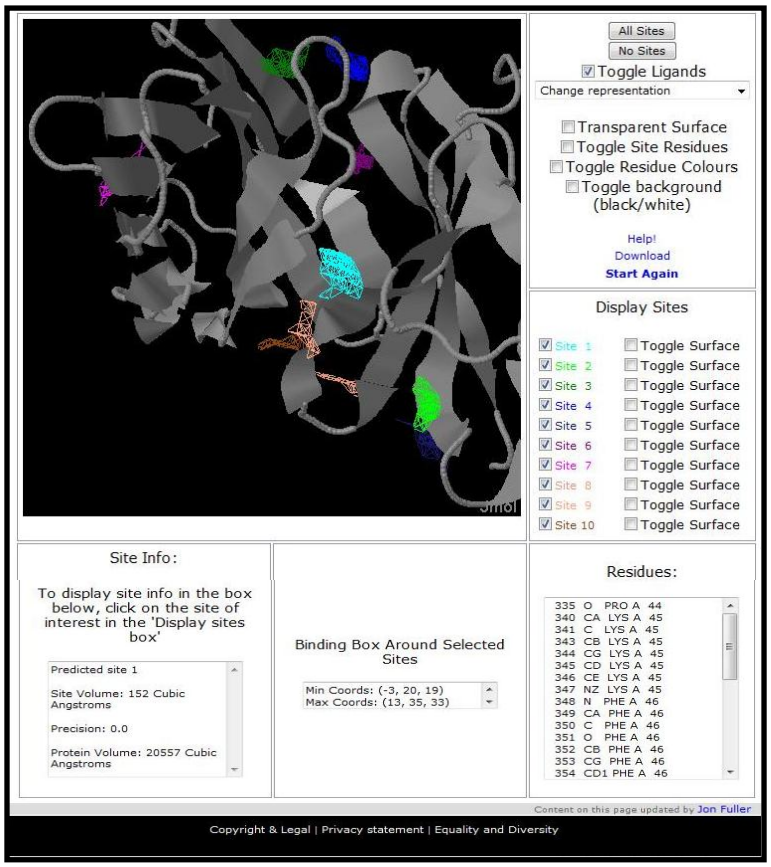


Figure 5. 1Screenshot of Q-site finder

5.2 Protein preparation of molecular targets of GBM

Protein-ligand docking was conducted on three dimensional (3D) structures of EGFRvIII molecular target of GBM (PDB ID 1I8I). This structures was obtained from protein data bank[17]. Prior to the docking, hydrogen atoms were added to the crystal structure of 1I8I and the structure was minimized by addition of required Gasteiger charges using AutoDock4.2v.All the missing loops were filled. Charge field was set up and extra Hydrogen atom was added and all unwanted water molecules was removed.

Total Gasteir charges added = 0.002

Total Kollman charges added = -67.7

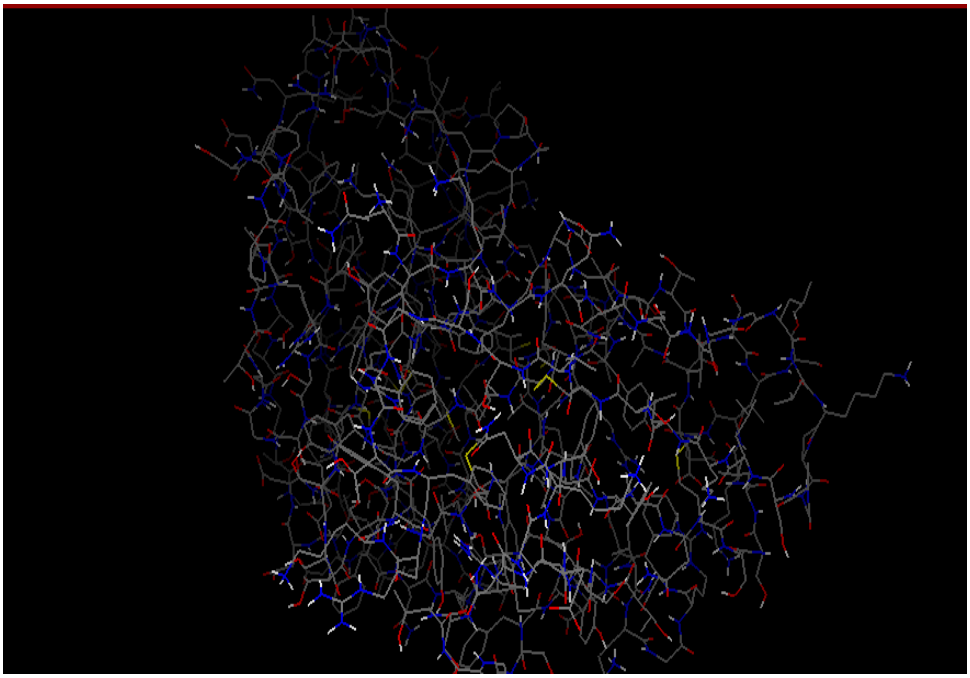


Figure 5. 2Minimized structure of EGFRvIII prior to docking using Autodock

5.3 Docking Using AutoDock 4.2

As mentioned earlier, docking with AutoDock follows three important steps :

1. Preparation of receptor & ligand files.

2. Calculation of affinity maps by using a 3D grid around the receptor & ligand.

3. Defining the docking parameters and running the docking simulation.

The preparation step starts with pdb files of receptor (R.pdb) and ligand (L.pdb), which are added with hydrogen and then saved as RH.pdb & LH.pdb. The calculation of affinity maps in the "Grid" section requires the above pdb files to be assigned charges & atom types, and also that the nonpolar hydrogens are merged. This is done automatically by ADT, and the resulting files are saved as RH.pdbqt & LH.pdbqt, which is the only format AutoGrid & AutoDock can work with. Calculation of affinity maps is done by AutoGrid, and then

docking can be done by AutoDock. The newest docking algorithm is LGA (Lamarckian Genetic Algorithm).

Step	Input	Output	File contains:
Preparation of receptor and ligand files	R.pdb	R_rigid.pdbqt	Rigid part of the receptor
		R_flex.pdbqt	Flexible part of the receptor
	L.pdb	L.pdbqt	
AutoGrid	R_rigid.pdbqt		
	R.gpf		Grid parameters
		R.glg	Grid log file (not used)
		R.*.map	Atom-specific affinity maps
		R.maps.fld	Grid_data_file
		R.d.map	Desolvation map
		R.e.map	Electrostatic map
Preparation for docking	R_rigid.pdbqt	L.pdf	Docking parameters
	R_flex.pdbqt		
	L.pdbqt		
Docking	R_rigid.pdbqt	L.dlg	Log + coordinates + energies
	R_flex.pdbqt		
	L.pdbqt		
	L.pdf		
	R.*.map		
	R.maps.fld		
	R.d.map		
	R.e.map		

Figure 5. 3 Screenshot of the files during docking process

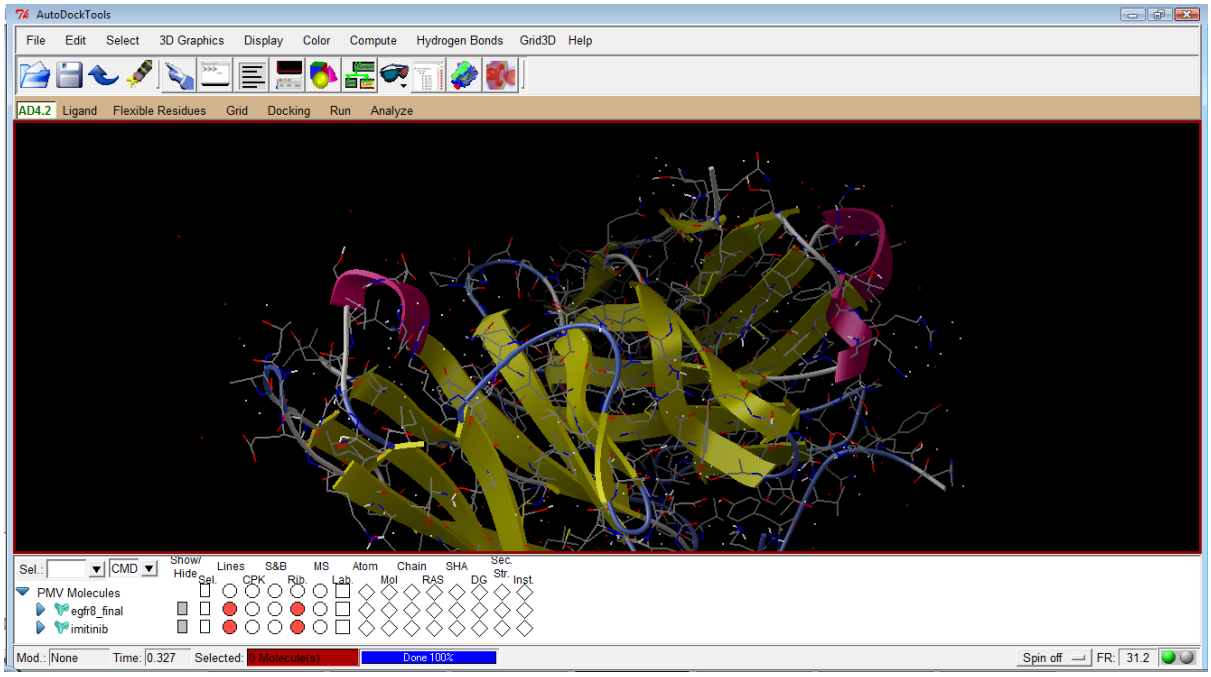


Figure 5. 4 ScreenShot during Docking using AutoDock4.2

5.4 Modification of ligands structure

The three ligands forming the most stable complexes with minimum binding energies are choosen for further analysis. The structure of the ligands were modified in an attempt to derive better inhibitor to the receptor. Some of the alterations considered were tautomeric transforms and bioisotertic substitutions.

Tautomeric transforms of the ligand were checked using Chemsketch 12.0.

Bioisosteric substitution is the process of starting withand intital active compound and replacing functional groups with other groups that have similar biological properties. Unlike isosteres, which have the same number and arrangement of electrons, bioisosteres are selected based on having a similar electronegativity, steric size , lipophilicity and other properties that affect the way in which drug binds to the target.

Monovalent Groups : H, F , Cl, Br, I, OH, SH, NH_2, SCH_3, PH_2, CH_3, i-Pr, t-Bu, OCH_3.

CHAPTER 6

RESULTS AND DISCUSSION

6. RESULTS AND DISCUSSIONS

6.1 MINIMIZATION OF PROTEIN

First step prior to docking process is the minimization of the protein. So after the PDB file of 1I8I (EGFRvIII) was downloaded , its minimization was done as follows:

Addition of Hydrogen Atoms = 0

Addition of Kollman Charges = -67.7

Addition of Gasteiger Charges =0.0002

Following figure shows the molecule before and after minimization was performed.

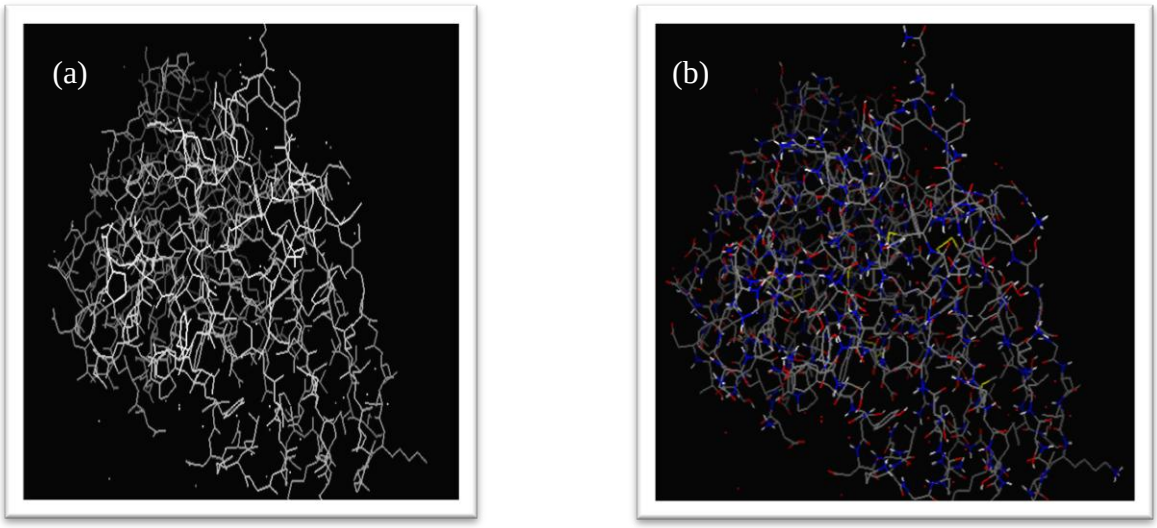


Figure 6. 1 EGFRvIII (a) before minimization (b) after minization

6.2 PROTEIN BINDING SITE

Protein binding site prediction using Q-Site Finder, gave 10 probable sites of ligand binding. These are the site which might be used by the docking programs before providing the best docking site and the corresponding result.

The Q-SiteFinder predicted the following:

Predicted site: 1

Site Volume; 152 Cubic

Protein volume: 20557 cubicA

Binding box around selected site:

Min Coordinate: (-3, 20, 19)

Max Coordinate.: (13, 35, 33)

Therefore, Central coordinates to be used as a input in the docking studies is (5, 27.5, 21)

All of the protein atoms close to a probe-cluster defining a site are listed below. The atom number, atom type, residue name, chain identifier and residue number are given

340 CA LYS A 45

341 C LYS A 45

343 CB LYS A 45

344 CG LYS A 45

345 CD LYS A 45

346 CE LYS A 45

347 NZ LYS A 45

348 N PHE A 46

349 CA PHE A 46

350 C PHE A 46

351 O PHE A 46

352 CB PHE A 46

353 CG PHE A 46

354 CD1 PHE A 46

356 CE1 PHE A 46

418 CA ARG A 55

419 C ARG A 55

420 O ARG A 55

421 CB ARG A 55

422 CG ARG A 55

423 CD ARG A 55

424 NE ARG A 55

425 CZ ARG A 55

426 NH1 ARG A 55

428 N PRO A 56

429 CA PRO A 56

430 C PRO A 56

431 O PRO A 56

432 CB PRO A 56

433 CG PRO A 56

434 CD PRO A 56

439 N VAL A 58

443 CB VAL A 58

445 CG2 VAL A 58

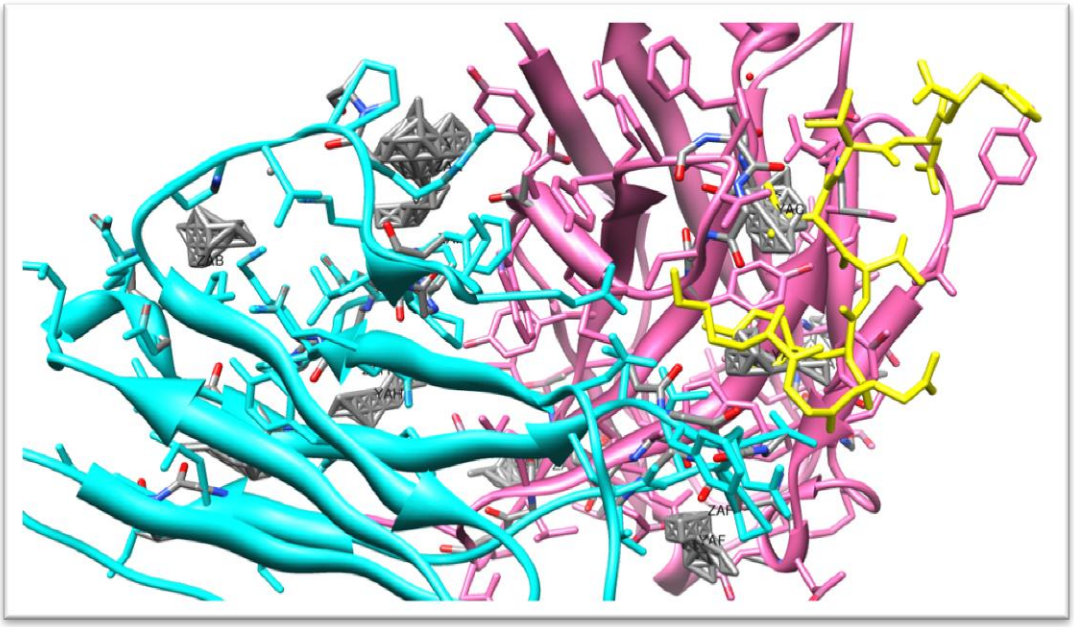
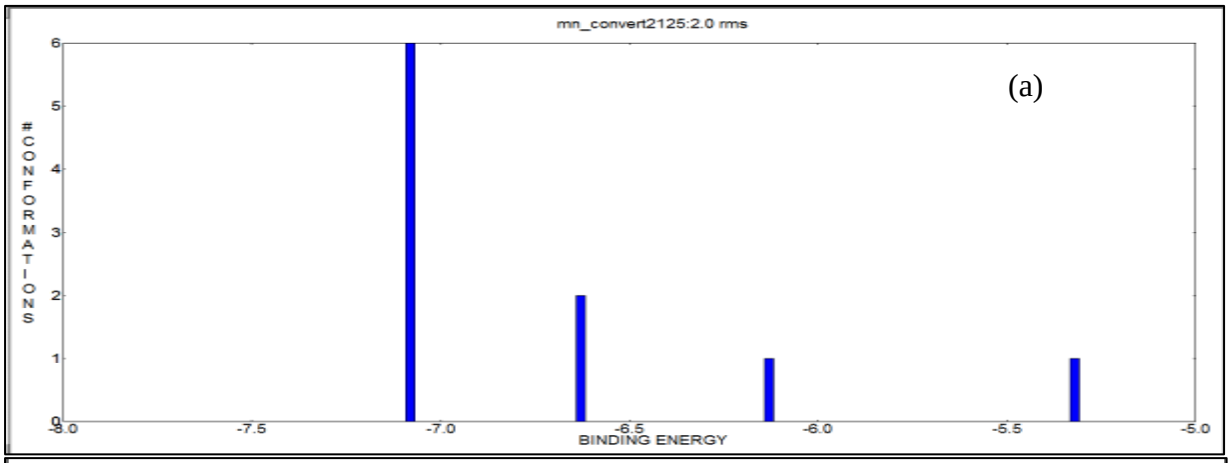
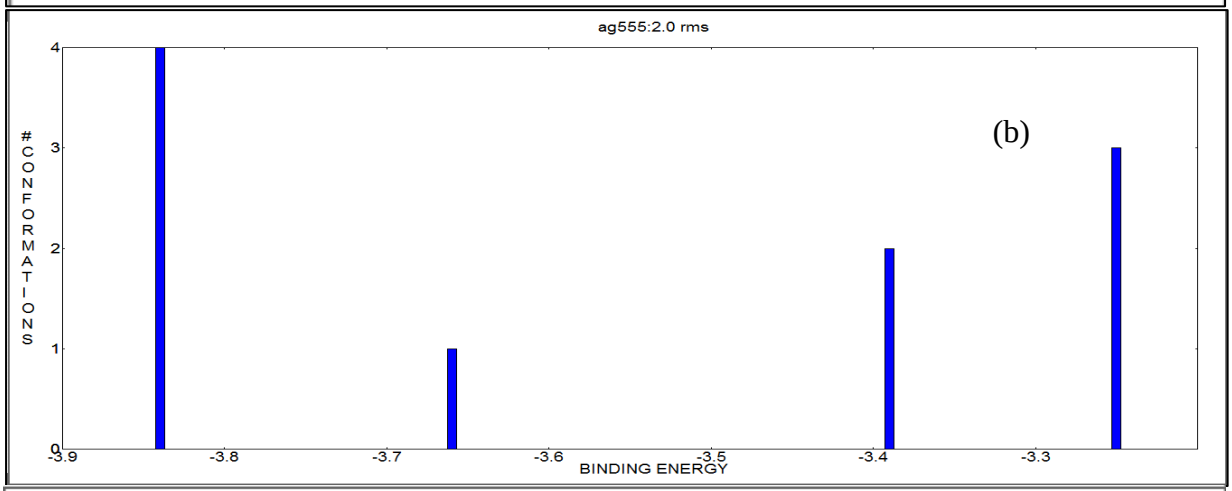


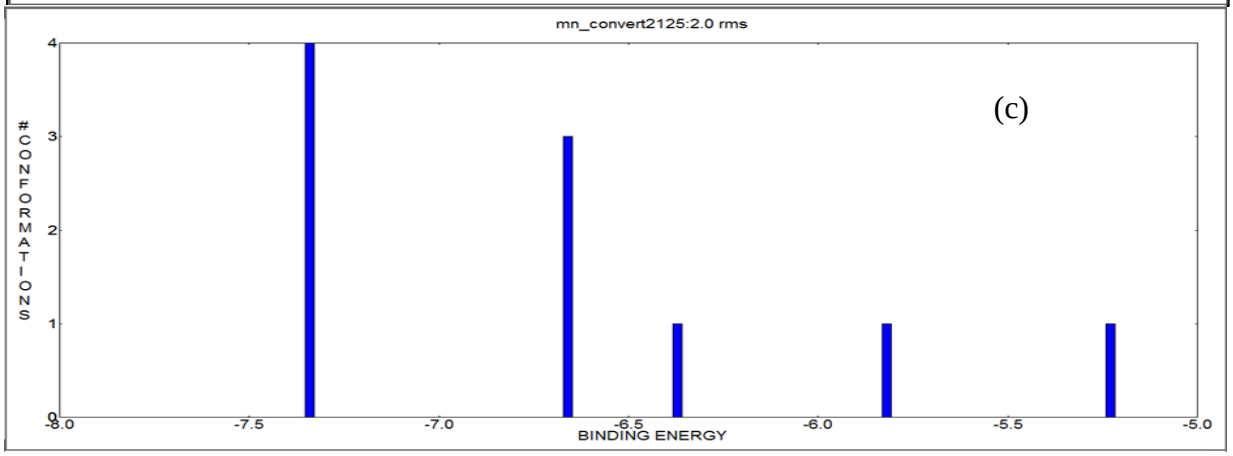
Figure 6. 2 Probable binding site of EGFRvII using Q-Site Finder

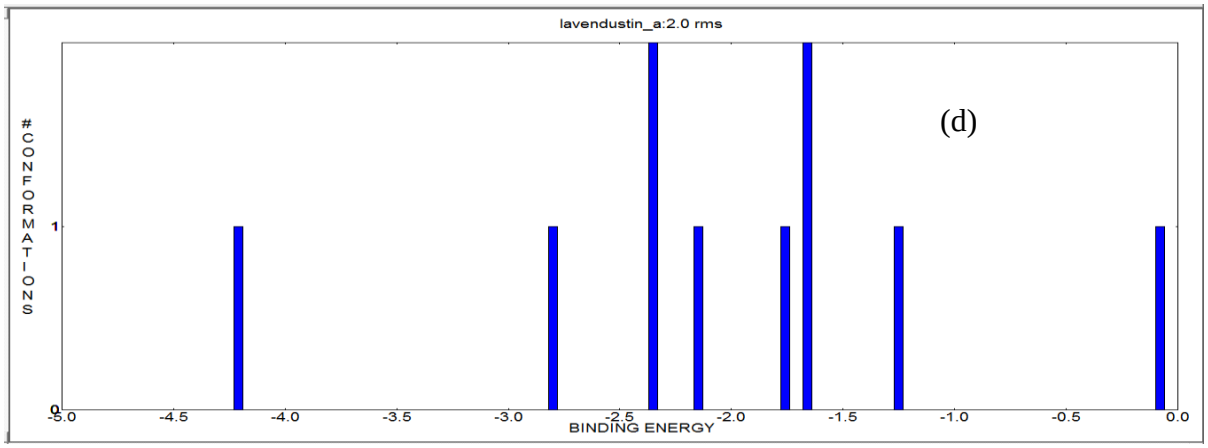
6.3 RESULTS OF DOCKING IN AUTODOCK4.2

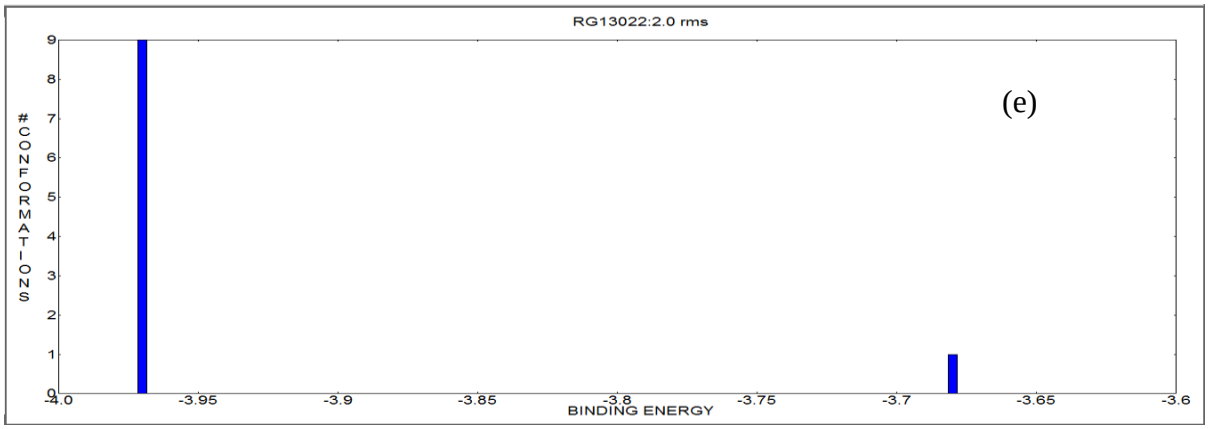
After the docking was performed using AutoDock4.2 which calculated the free binding energy values for each protein-ligand complex formed, the stability of the inhibitor-receptor binding was studied on the basis of relative low binding energy of the complex. Below shown are the conformational clustering histograms for all the protein-ligand complex.

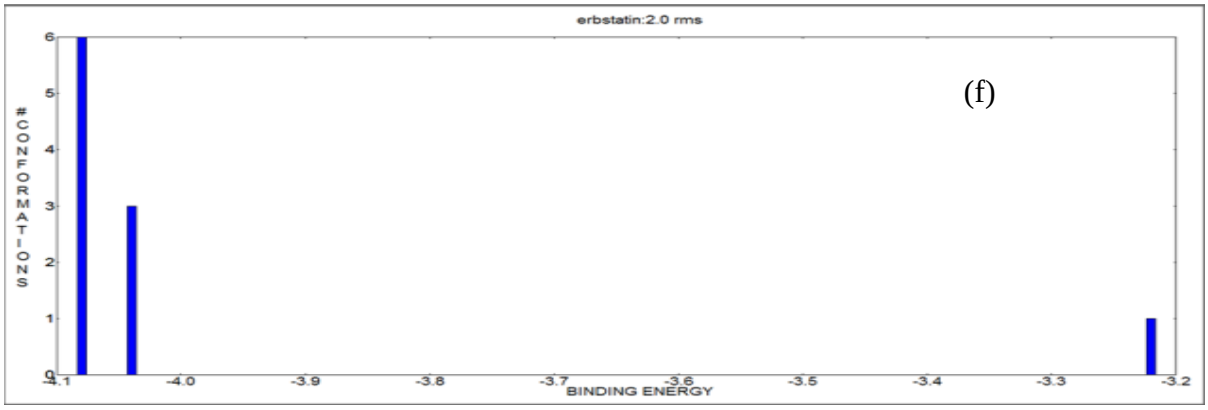


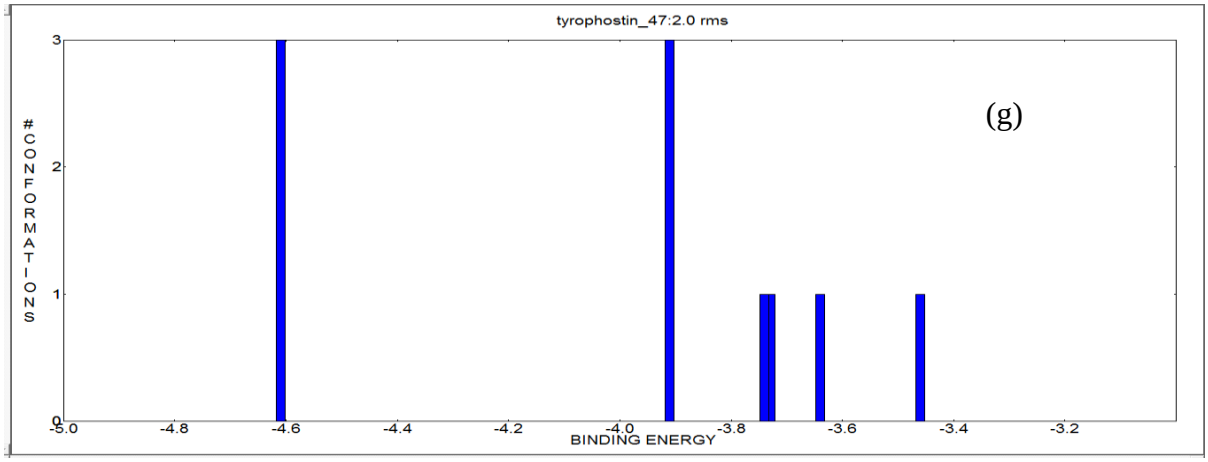


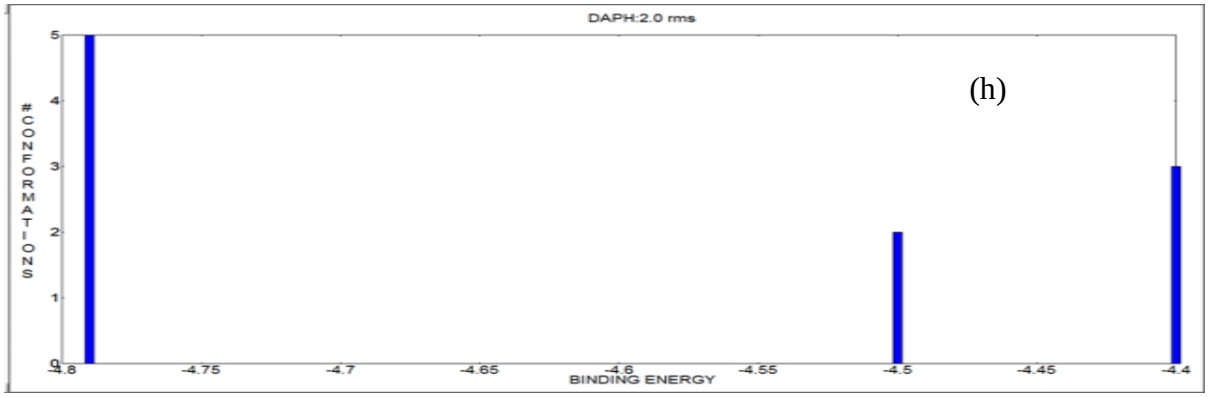


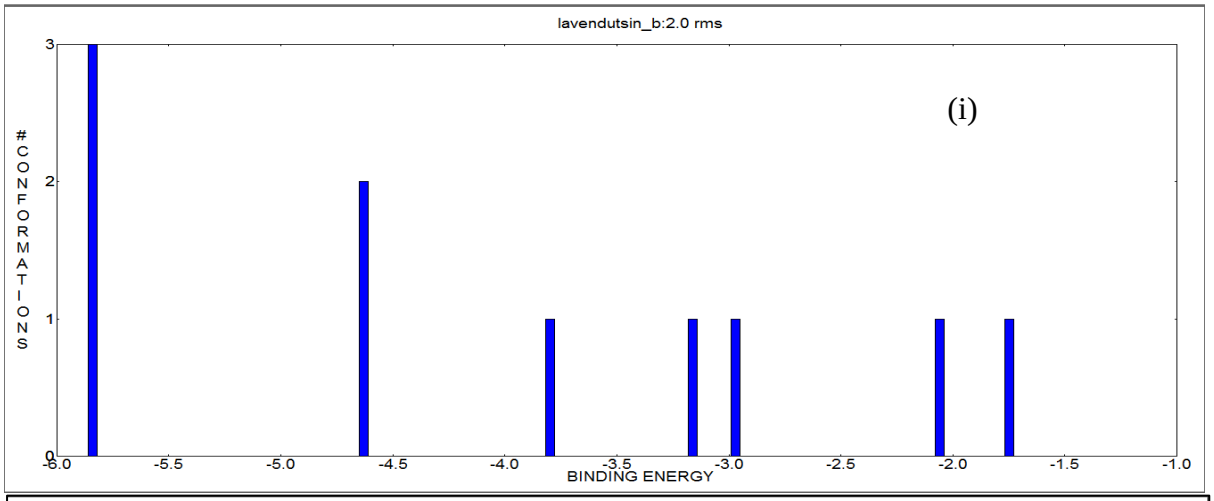

lavendustin_a:2.0 rms
#CONFORMATIONS
BINDING ENERGY
(d)

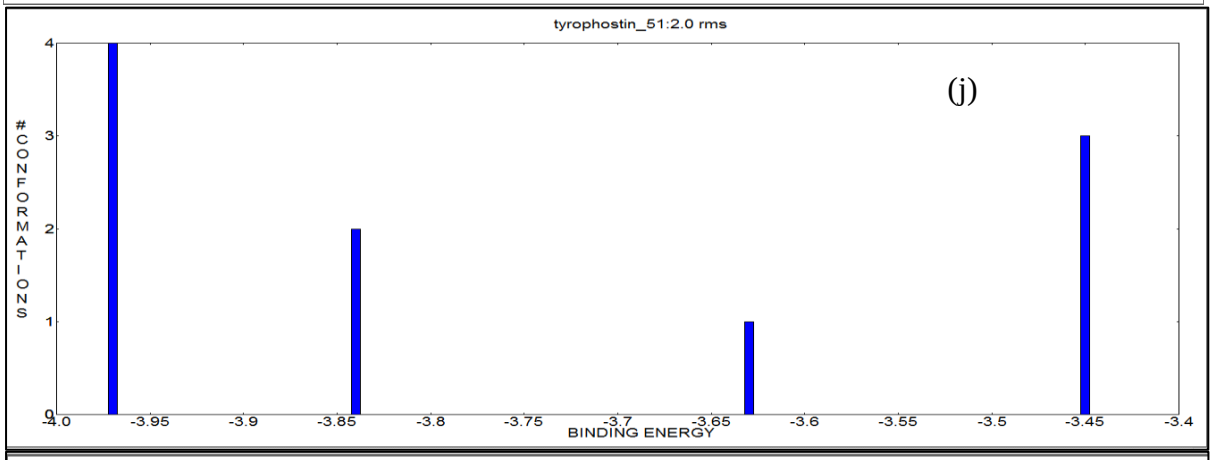

RG13022:2.0 rms
#CONFORMATIONS
BINDING ENERGY
(e)

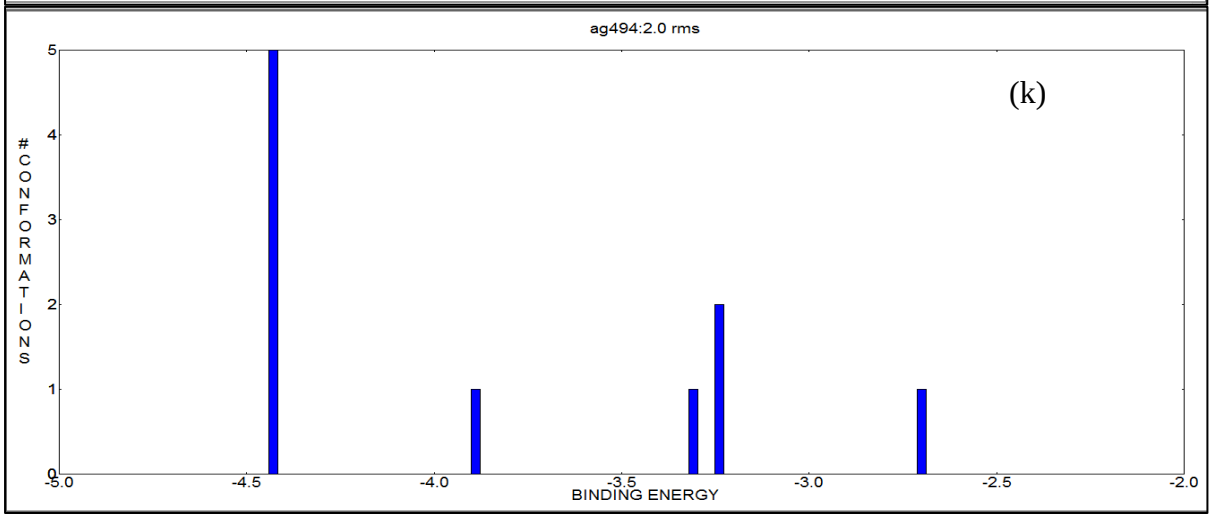

erbstatin:2.0 rms
#CONFORMATIONS
BINDING ENERGY
(f)

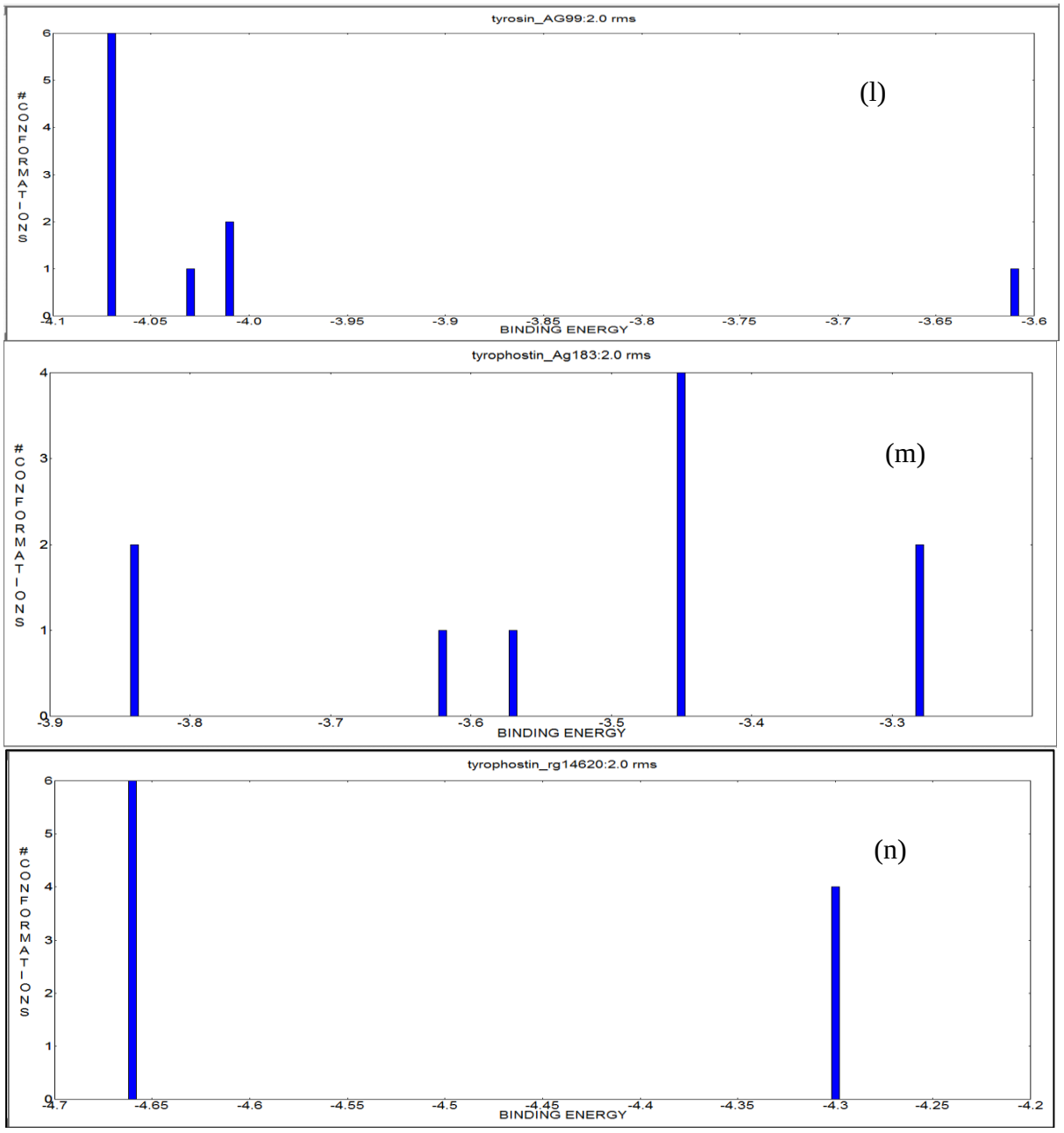

tyrophostin_47:2.0 rms
#CONFORMATIONS
BINDING ENERGY
(g)


DAPH:2.0 rms
(h)
#CONFORMATIONS
BINDING ENERGY


lavendutsin_b:2.0 rms
(i)
#CONFORMATIONS
BINDING ENERGY


tyrophostin_51:2.0 rms
(j)
#CONFORMATIONS
BINDING ENERGY


ag494:2.0 rms
(k)
#CONFORMATIONS
BINDING ENERGY


tyrosin_AG99:2.0 rms
#CONFORMATIONS
BINDING ENERGY
(l)
tyrophostin_Ag183:2.0 rms
#CONFORMATIONS
BINDING ENERGY
(m)
tyrophostin_rg14620:2.0 rms
#CONFORMATIONS
BINDING ENERGY
(n)

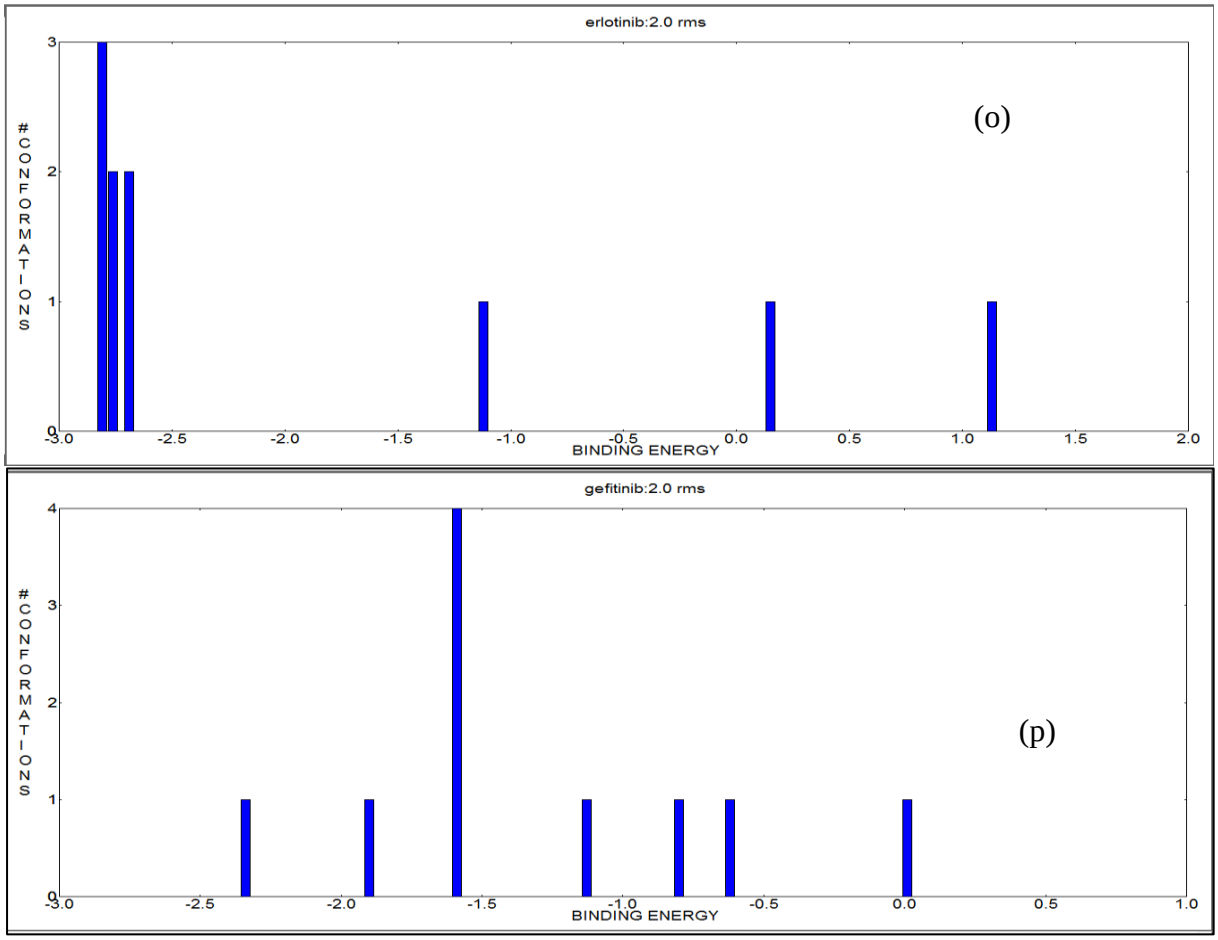


Figure 6. 3 Binding Energy Plots of (a)PD153035 (b)AG555 (c)BIBX-1328BS (d) lavendustin A (e) RG13022 (f) erbstatinanalog (g)tyrophostin 47 (h)DAPH (i) Lavendustin B (j)tyrophostin 51 (k) AG494 (l)tyrophostin AG99 (m) tyrophostin AG183 (n) tyrophostin RG14620 (o) erlotinib (p)gefitinib

Table 6. 1 Free Binding Energy values of the Docked Complexes and the corrsponding Ligand used

LIGAND	FREE BINDING ENRGY IN Kcal/Mol
AG494	-4.77
TYROPHOSTIN AG999	-5.11
ERBSTAIN ANALOG	-5.58
TYROPHOSTIN 47	-4.93
RG 13022	-4.63
TYROPHOSTIN 183	-5.19
TYROPHOSTIN 51	-5.06

DAPH	-4.77
LAVENDUSTIN B	-6.56
LAVENDUSTIN A	-4.41
BIBX-1382BS	-7.9
PD 153035	-8.98
ERLOTINIB	-5.15
GEFITINIB	-4.52

For a precise comparative study of the free binding energies of the complexes formed , the above result was plotted in a graph.

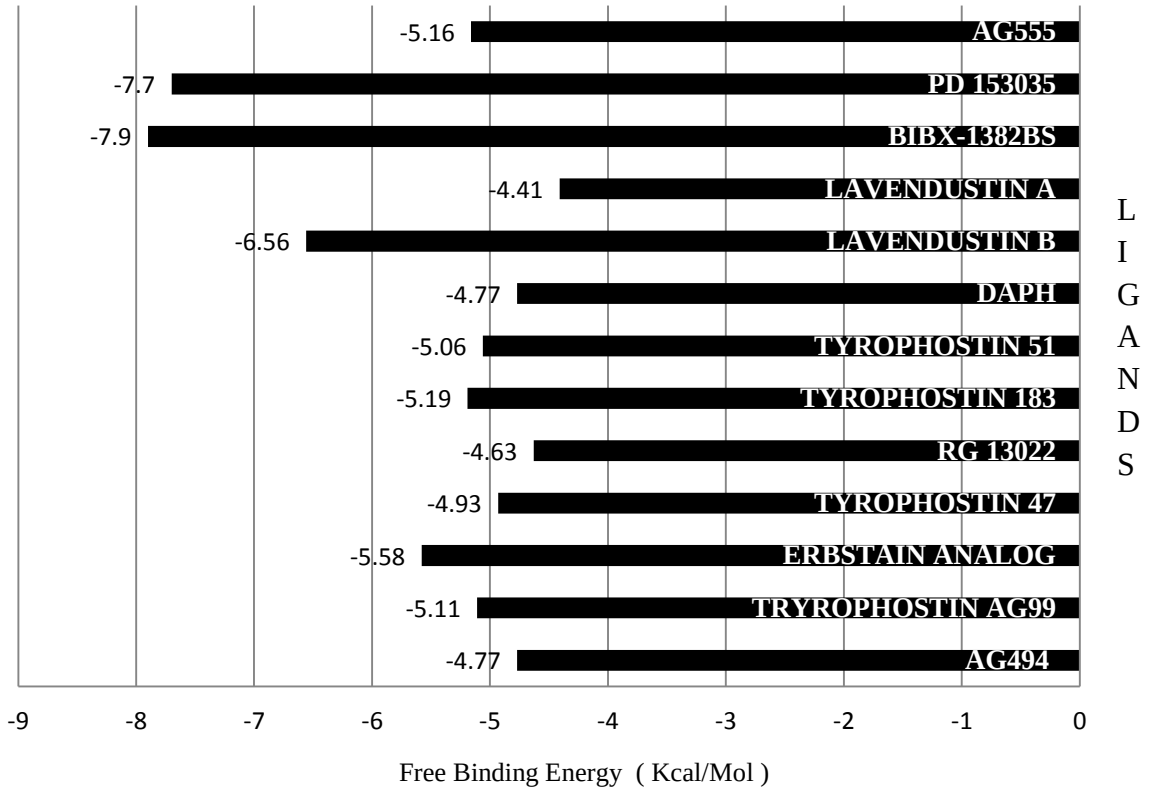


Figure 6. 4 Graph showing the comparative Free Binding energy plots of the ligands

By comparing all the free binding energies in case of each ligand , the most favourable ligands forming the docked structures were :

- PD153035
- LAVENDUSTIN A
- BIBX -1382BS

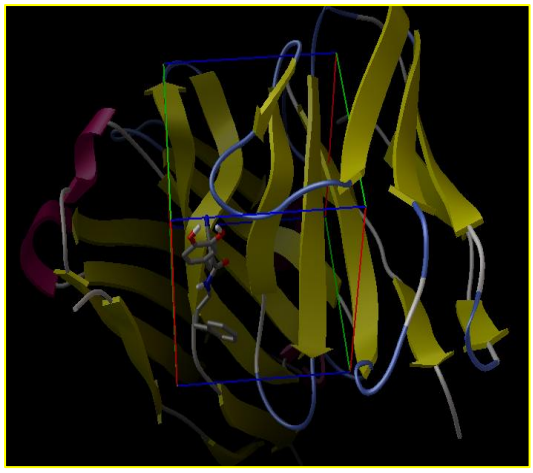


Figure 6. 5 . Docked complex with PD153035

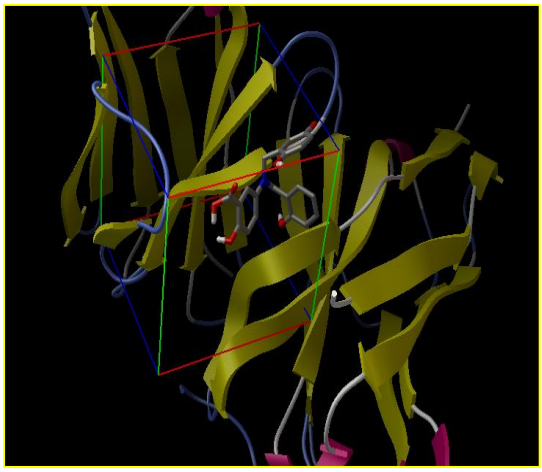


Figure 6. 6 Docked Complex with Lavendustin A

Figure 15. Docked Complex with Lavendustin A

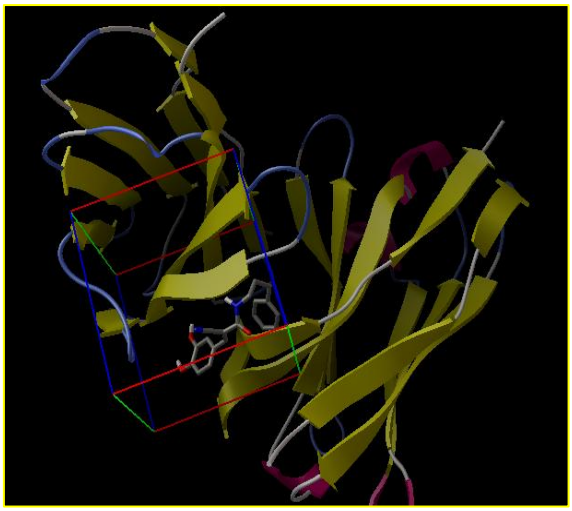


Figure 6. 7 . docked Complex formed with BIBX-1328BS

6.6 MODIFICATION OF THE LIGAND STRUCTURE

Several modifications were done on the three most stable ligands which were obtained from docking with 1I8I using HEX6.3. ChemSketch was used to modify the structures. It was on basis of finding tautomeric structures and performing biosteric substitutions.

However, Chemsketch showed that the molecules were already in its most stable tautomeric form. Hence, monovalent groups of the structure were changed in an attempt to derive a structure capable of providing a more stable docking with 1I8I.

Modifications

PD153035

Some of the monovalent Bioisters are H, F, Cl, Br, I , -OH, -SH, $-NH_2$, $-SCH_3$, $-PH_2$, $-CH_3$, i-Pr, t-Bu . All these groups contains similar electronegativity, steric size ,lipophilicity and other properties that affects the way in which drug binds to the target.

The bromide monovalent group was substituted with *t-butyl* group and the following structure was obtained. let's say M1.

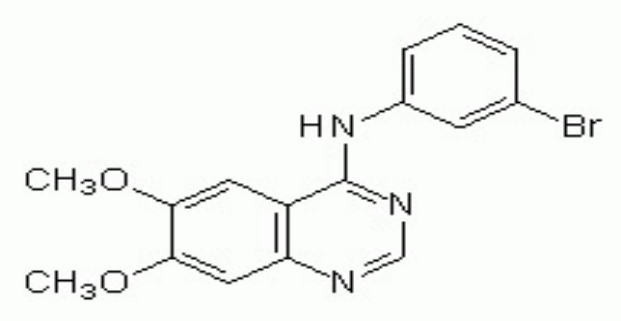


Figure 6. 8 PD15303 . Source PDB

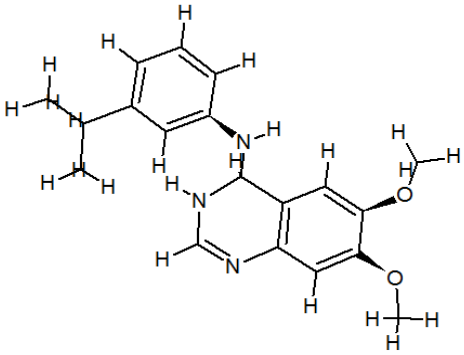


Figure 6. 9 Modified structure, drawn using ChemSketch. M1

Further one of the ether group (-OCH_3) was substituted with methyl group, lets say M2

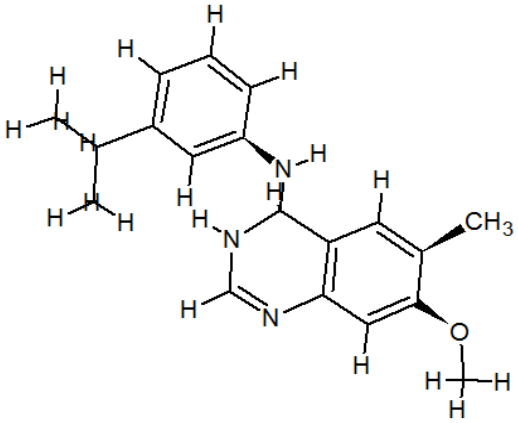


Figure 6. 10 Modified structure drawn using ChemSketch. M2

BIBX-1328BS

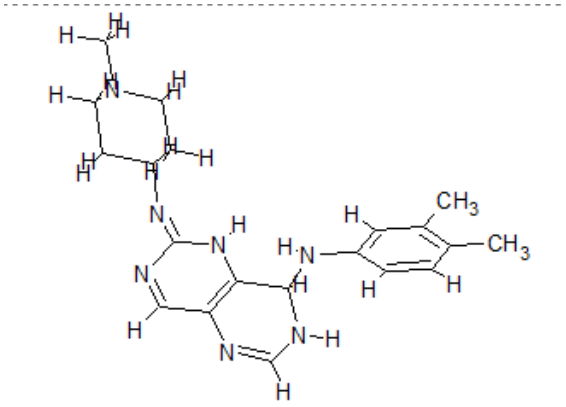


Figure 6. 11 modified structure using ChemSketch. (B1)

6.7 DOCKING WITH THE MODIFIED STRUCTURES

The free binding energy value calculated after docking of M1 and M2 with 1I8I using Auto Dock was calculated to be -8.39Kcal/Mol and – 8 .21Kcal/mol . This shows that, these structural modifications might share similar properties as PD153035.

However , when this particular structure was docked , the binding energy value was not as favourable as in the previous case. Results showed it to be -4.1 Kcal/Mol

CHAPTER 7

CONCLUSION

7. CONCLUSION

By performing *in-silico* protein-ligand docking analysis of 16 non-specific inhibitors of EGFR with EGFRvIII receptors, this was an attempt to propose the most suitable commercially available EGFR inhibitors which can prove to be effective for targeting EGFRvIII mutant. Results showed that PD 153035, BIBX-1382BS and LAVENDUSTIN B formed protein-ligand complex with the least free binding energy, showcasing higher stability of the molecule formed as compared to the rest of the ligands. It was in fact observed that these ligand formed more stable complexes with the specific EGFRvIII mutant in comparison to Gefitinib and Erlotininb - the widely used EGFR inhibitors for molecular targeted cancer therapy. Further, biosteric substitutions were performed on these favourable inhibitors for structural modifications and in-silico protein-ligand docking was repeated with the modified structure. complexes showed free binding energy in the favourable range, leadin us to conclude that these few structures were could be proposed for further wetlab analysis inorder to study there detailed effect as molecular targeted drug therapy for treating Glioblastoma.

References

REFERENCES

1. Nagane M, Pan G, Weddle JJ, Dixit VM, Cavenee WK, Huang HJ. (2000).*Cancer Res* **60**: 847–853
2. Lou E. *Ann Med.* 2004 **36**: 2 [PMID: 15000342]
3. Walker C *et al*. *Clin Cancer Res.* 2004 **10**: 7182 [PMID: 15534091]
4. Oyasu R *et al. Cancer Res.* 1970 **30**: 1248 [PMID: 5426930]
5. Lipsitz D *et al. Vet Pathol.* 2003 **40**: 659 [PMID: 14608019]
6. *Schlegel J et al.,"**Amplification and differential expression of members of the erbB-gene family in human glioblastoma**",* J. Neurooncol *(1994);22(3): pp. 201–207*
7. *Schlegel J et al.,"*Amplification and differential expression of members of the erbB-gene family in human glioblastoma*",* J. Neurooncol *(1994);22(3): pp. 201–207.*
8. Heimberger AB, Hlatky R, Suki D, Yang D, Weinberg J, Gilbert M, Sawaya R, Aldape K:**Prognostic effect of epidermal growth factor receptor and EGFRvIII in glioblastoma multiforme patients.** *Clin Cancer Reviews* 2005, **11:**1462-1466.
9. Buckner, R. L. & Carroll, D. C. 2007 Self-projection and the brain. Trends Cogn. Sci. 11, 49–57
10. Stupp R., Hegi M.E., et al.,"**effects of radiotherapy with concomitant and adjuvant temozolomide versus radiotherapy alone on survival in Glioblastoma**"-5 years analysis of EORTC-NC (2009)
11. Burger P.C. and Vogel F.S. (1982). Surgical pathology of the nervous system and its coverings. 2nd ed. John Wiley and & Sons. New York. pp 226-266
12. Hoelzinger, D. B., Demuth, T., and Berens, M. E. (2007). Autocrine factors that sustain glioma invasion and paracrine biology in the brain microenvironment. J. Natl. Cancer Inst. 99, 1583–1593.

13. James CD, Olson JJ. Molecular genetics and molecular biology advances in brain tumours .*Currr Ooin Oncol,*1995 ;8: 188-195
14. Scherer H J*, "***The forms of growth in gliomas and their practical significance***",* Brain *(1940);63(part 1): pp. 1–35.*
15. Van Wijngaarden, Edwin; Dosemeci, Mustafa (2006). "**Brain cancer mortality and potential occupational exposure** to lead: Findings from the National Longitudinal Mortality Study, 1979–1989". *International Journal of Cancer* **119** (5): 1136–44
16. Stupp, Roger; Mason, Warren P.; Van Den Bent, Martin J.; Weller, Michael; Fisher, Barbara; Taphoorn, Martin J.B.; Belanger, Karl; Brandes, Alba A. et al (2005). "**Radiotherapy plus Concomitant and Adjuvant Temozolomide for Glioblastoma**". *New England Journal of Medicine***352** (10): 987–996
17. Berman HM *et al. Nucleic Acids Res.* 2000 **28**: 235 [PMID: 10592235]
18. Schrödinger, version 5.5, Schrödinger, LLC 2009
19. Michel F. Sanner. Python: A Programming Language for Software Integration and Development. J. Mol. Graphics Mod., 1999, Vol 17, February. pp57-61
20. Michel F. Sanner. Python: A Programming Language for Software Integration and Development. J. Mol. Graphics Mod., 1999, Vol 17, February. pp57-61